FIG. 1A

```
  1  GGCACGAGGC AAGCTTCCA  GGTTATCGTG ACGCACCTTG AAAGTCTGAG AGTACTGCC  CTACAGAAAG TTACTAGTGC CCTAAAGCTG GCGCTGGCAC
     CGGTGCTCCG TTCGGAGAGT CCAATAGCAC TGCGTGGAAC TTTCAGACTC TCGATGACGG GATGTCTTTC AATGATCACG GGATTCGGAC CGCGACCGTG
     ^start insert

101  TGAYGTYACT GCTGCTGTTG GAGTACNACT TCCCTATAGA AAACACTGC  CAGCACTTA  AGACACTCA  CACTTCAGA  GTGGCCTTGA GAAGATTTG
     ACTACAATGA CGACGACAAC CTCATGYTGA AGGGATATCT TTTGTGTGACG GTCGTGAGT  TCTGTGTAGT GTGGAAGTCT CACCGGAACT CTTCTAAAC
     ^Met
  1   M  L   L   L   E   Y  N  F  F   P  I  E   N  N  C   Q  H  L   K   T  T  T   H  T  F  R   ^possible splice donor

201  GGGTCAAGGA TCATACGACGA GAACACCACT TAAGAGCATA GTGAACTAGT CTGCGATGTGA GACGCTGAGA TCCTATATCA GGCTGTGATA GGAGGGAAAC
     CCCAGTTCCT AGTACTGCGT CTTGTGGTGA ATTCTCGTAT CACTTGATCA GACGCTACACT CTGCGACTCT AGGATACAGT CCGAACACTAT CCTCCCTTTG
  1   V   K   N   L   N   P   K   F   S   I   H   D   Q   D   H   K   V   L   V   L   D   S

301  AGAAACCAAA GGAAGAACA  CCCTTAAGA  GCCTTAGGA  GGTTAAGGA  CATTCYTGAC AGTCACYGGC CCACCCTGGG GGCCCGTGTT CTTTATCTAAA
     TCTTTGGTTT CCTTCTTGT  GGGAATTCT  CGGAATTCCTAT CCAATTCCTAT GTAAGAACTG TCAGTGGACCG GGTGGGACCC CCGGGACACA GAAATAGTTT

401  CAAGGTCCTG AGCTGCTTGC AGAAGTCGCAA AGGTGAAGAA CTTAAACCCG AAGAAATTCA GCAAGCATCA CCAAGATCAC AAAGTACTGG TCCTCGACTC
     GTTCCAGGAC TCGACGAACG TCTCCAGGTT TCCACTTCTT GAATTTGGGC TTCTTTAAGT CGTTCGTAGT GGTTCTAGTG TTTCATGACC AGGAGCTGAG
     ^possible splice acceptor
  1   V   K   N   L   N   P   K   F   S   I   H   D   Q   D   H   K   V   L   V   L   D   S

                    85066.AH1282.Asc.f., 5'Tag: TTTCCCTTGGCGCGCC
                    85066.AH1284.Nsi.f., 5'Tag: TTTCCCTTATGCATCAGATGACGATGACAAA
                    ^possible splice acceptor
                    ^ORF

501  TGGGAANTTC ATAGCAGNTC CAGATAAAAA CTACATACCGC CCAGAGAATCT TCTTTGCATT AGCCTCANTCC TTGAGCTCAG CCCTCGCCGA GAAAGGAAGT
     ACCCTTAAGA TATCGTCAAG GTCTATTTTT GATGTATGCGG GGTCTCTTAGA AGAAACGTAA TCGGAGTNAGG AACTCGGATCTC GGGAGCGGCT CTTTCCTTCA
 24   G   N   L   I   A   V   P   D   K   N   Y   I   R   P   E   I   F   F   A   L   A   S   S   L   S   S   A   S   A   E   K   G   S

601  CCGAATCCTC TGGGGAGTTC TAAAGGGGAG TTTGGTGTCT ACTGTGACAA GGATAAAGGA CAAAGTCATC CATCCCTTCA GCTTAAGAAG GAGAAACTGA
     GGCTTAGGAG ACCCCGAGAG ATTTCCCCTC AAACAGAGA  TGACACTGTT CCTATTTCCT GTTTCAGTAG GTAGGGAAGT CGACTCGTTC CTCTTTGACT
 57   P   I   L   L   G   V   S   K   G   E   F   C   L   Y   C   D   K   D   K   G   Q   S   H   P   S   L   Q   L   K   K   E   K   L   M
```

```
 701 TGAAGCTGGC TGCCCAAAAG GAATCAGCAC GCCGGCCCTT CATCTTTTAT AGGGCTCAGG TGGGCTCCTG GAACAATGCTG GAGTCGGACGG CTCACCCCGG
  91 ACTTCGACCG ACGGGTTTC CTTAGTCCTG CGGCCGGGAA GTAGAAATA TCCCGAGTCC ACCGGAGGAC CTTGGTAGGAC CTCAGCCGCC GAGTGGGGCC
     K   L   A   A   Q   K   E   S   A   R   P   F   I   F   Y   R   A   Q   V   G   S   W   N   M   L   E   S   A   H   P   G

 801 ATGGTTCATC TGCACCTCCT GCCAATTGTA TGAAGCCATT GGGGTGACAG ATAAATTGA GAACAGGAAA CACATTGAAT TTTCATTTCA ACCAGTTTGC
 124 TACCAAGTAG ACGTGGAGGA CGTTACACAT ACTTCGGACAA CCCCACTGTC TATTTAAACT CTTGTCCTTT GTGTAACTTA AAAGTAAAGT TGGTTCAAACG
     W   F   I   C   T   S   C   N   C   N   E   P   V   G   V   T   D   K   F   E   N   R   K   H   I   E   F   S   F   Q   P   V   C

 901 AAAGCTGAAA TGAGCCCCAG TGAGGTCAGC GAATTAGGAAA CTGCCCCCATT GAACGCCTTC CTGGCTTATT TGAACTTAATT GTATAAAAAC ACCAAAACTG
     TTTCGACTTT ACTCGGGGTC ACTCAGTCG CTTATCCTTT GACGGGGTTAA CTTGCGGAAG GACGCGATTAA ACTTGATTAA CATATTTTTG TGGTTTTGAC
     K   A   E   M   S   P   S   E   V   S   D   O

 157                                              ^85066.AH1283.r
                                                  ^85066.AH1285.Not.r, 5'Tag: TTTCCCTTTGCCGCCCGCTTA

1001 CTCACT
     GAGTGA
```

Fig 1B

```
  1  TAATTCACCA TGTCTGCACT TGTTATCCTA GCTCTCGTTG GAGCTCCAGT TCTGACTTAC AAAGACGATG ACGACAAGCT TGCGGCCGCG AATTCAGCTC
     ATTAAGTGGT ACAGACGTGA ACAATAGGAT CGAGAGCAAC CTCGAGGTCA AGACTGAATG TTTCTGCTAC TGCTGTTCGA ACGCCGGCGC TTAAGTCGAG
  1     M  S  A  L  I  L  A  L  V  G  A  A  V  A  D  Y  K  D  D  D  D  K  L  A  A  A  N  S  A  L
        ^orf                                            ^flag                          ^insert st

101  AGTTCCAGAT AAAAAACTGA TAGGCCAGAG ACTTCTTTGC ATTAGCCTCA CCTTGAGCTC AGCCTCTGCG GAGAAAGGAA GTCCGATTCT CCTGGGA
     TCAAGGTCTA TTTTTTGACT ATCCGGTCTC TGAAGAAACG TAATCGGAGT GGAACTCGAG TCGGAGACGC CTCTTTCCTT CAGGCTAAGA GGACCCT
 32   C  R  G   P  K  V   K  N  L  N   P  K  K   F  S  I   H  D  Q  D   H  K  V  L   V  L   D  S  G  N   L  I  A

201  AGTTCCAGAT TACGGCCAGA GATCTCGTTT CACCTTGAGC ATCTTGAGCT CGAGGAGACG TAGAGACTCG AGTCGGAGAC GCCTCTTTCC TTCAGGCTA
     TCAAGGTCTA ATGCCGGTCT CTAGAGCAAA GTGGAACTCG TAGAACTCGA GCTCCTCTGC ATCTCTGAGC TCAGCCTCTG CGGAGAAAGG AAGTCCGAT
 65   V  P  D   K  N  Y   I  R  P  E   I  F  F   A  L  A   S     S  A  S   A  E  K  G   S  P   I  L  G

301  GTCTGTAAGG GGAGATTTTG TCTCTACTCT GACAAGGATA AAGGACAAAG TCTCACATCC CTTCAGCTGA AGAAAGAGAA ACTGATGAAG CTGGCTGCAC
     CAGACATTCC CCTCTAAAAC AGAGATGAGA CTGTTCCTAT TTCCTGTTTC AGAGTGTAGG GAAGTCGACT TCTTTCTCTT TGACTACTTC GACCGACGTG
 98   V  S  K  G   E  F  C   L  Y  C   D  K  D  K   G  Q  S   H  P  S   L  Q  L  K   K  E  K   L  M  K   L  A  A  Q

401  ACAAGCGCGG CCCTTCGAGC TCAGGTGGAC TTTTATAGGG CAAATTCCGG AGTACACGCG AGACCTGTTT ACAGCTGCAC ACCCGGATGG GTGGTTCAC
     TGTTCGCGCC GGGAAGCTCG AGTCCACCTG AAAATATCCC GTTTAAGGCC TCATGTGCGC TCTGGACAAA TGTCGACGTG TGGGCCTACC CACCAAGTG
132   K  E  S   A  R  R   P  F  I  F   Y  R  A   Q  V  G   S  W  N  M   L  E  S   A  A  H   P  G  W  F   I  C  T

501  CTCTGCGAT CTGTAATGAGC CTGTTGGGGT GACAGATAAA TTTGAGAACA GGAAACACAT TGAATTTTCA TTTCAACCAG TTTGCAAAGC TGAAATGAGC
     GAGACGCTA GACATTACTCG GACAACCCCA CTGTCTATTT AAACTCTTGT CCTTTGTGTA ACTTAAAAGT AAAGTTGGTC AAACGTTTCG ACTTTACTCG
165   S  C  N   C  N  E  P   V  G  V   T  D  K   F  E  N  R   K  H  I   E  F  S   F  Q  P  V   C  K  A   E  M  S

601  CCCAGTAAGG GGGTACAGCAT CCAATGGTCA CCTAGGATCT GGATCCCGGG
     GGGTCACTCC CCCATGTCGTA GGTTACCAGT GGATCCTAGA CCTAGGGCCC
198   P  S  E   V  S  D  D
                ^inserts ends
```

Fig 2

^insert st   ^inserts ends

```
  1  GGCACGAGGC AAGCCTTCCA GGTTATCGTG ACGCACGTTG AAAGTCGTAG AGTACTACCC CTACAGAAGG TTACTAGTGC CCTAAAGCTG GGCTGCGCAC
     CCGTGCTCCG TTCGGAAGGT CCAATAGCAC TGCGTGCAAC TTTCAGCATC TCATGATGGG GATGTCTTCC AATGATCACG GGATTTCGAC CCGACGCGTG

101  TGATGTTGACT GCTGCTCTTG CACTGTTG GAGTACGAACT TCCGTTATGA AAACGACTGG CAGCACTGTCA CACTTCAGA GTGAAGAACT TCAACCCGAA
     ACTACAATGA CGACGAGAAC CTCATGTTGA AGGCAATATCT TTTGTTGACG GTCGTGGACTT TCTGGTGAGT CACTTCTTGA ATTGGGCGTT

  1  M   L   L   L    E   Y   N   F    P   I   E   N    N   C   Q   H    L   K   T   T    H   T   F   R    V   K   N   L    N   P   K
     Met

201  GAAATTCAGC ATTCATCACC AGGATCACAA AGTACTGGTC CTGAGACTCAT GGAATTCAT AGCAATTCCA GATAAAAACT ACATACCCCC AGAAGATCTTC
     CTTTAAGTCG TAAGTAGTGG TCCTAGTGTT TCATGACCAG GACTCTGAGTA CCTTAGAGTA TCGTTAAGGT CTATTTTTGA TGTATGGGGG TCTTCTAGAAG

 34  K   F   S   I    H   D   Q   D    H   K   V   L    V   L   D   S    G   N   L   I    A   V   P   D    K   N   Y   I    R   P   E   I   F

301  TTTGCATTAG CCTCATCCTT GAGCTCAGCC TCTGCGGAGA AAGGAAGTCC GATTCTCCTG GGGTCTCTA AAGCGGAGTT TGTCTCTAC TGTGACAAGG
     AAACGTAATC GGAGTAGGAA CTCGAGTCGG AGACGCCTCT TTCCTTCAGG CTAAGAGGAC CCCAGAGAT TTCCCGTCAA ACAGAGATG ACACTGTTCC

 67  F   A   L   A    S   S   L   S    S   A   S   A    E   K   G   S    P   I   L   L    G   V   S   K    G   E   F    C   L   Y   C   D   K   D

401  ATAAAGCAGA AAGTCACCA TCCCTCAGC TGAACAGAG GAACTCATG AACGTCCGC CCCAAAAGGA ATCAGCACGC CGGCCCTTCA TCTTTTATAG
     TATTTCGTCT TTCAGTGGT AGGGAAGTCG ACTTGTCTC CTTGAGTAC TTGCAGGCG GGGTTTTCCT TAGTCGTGCG GCCGGGAAGT AGAAAATATC

101  K   G   Q   S    H   P   S   L    Q   L   E   L    M   K   L   A    A   Q   K   E    S   A   R    R   P   F   I    F   Y   R

501  GGCTCAGTG GGCTCGTGA ACATCATCGA GTCAGCCGAT CACGCCCGAT CACTCCTGC AATTGTAATG AGCGTGTTGG GGTGACAGAT
     CCGAGTCAC CCGAGCACT TGTAGTAGCT TCAGTCGGCTA GTGCGGGCCTA GTGAGGACG TTAACATTAC TCGGCACAACC CCACTGTCTA

134  A   Q   V   G    S   W   N   M    L   E   S   A    A   H   P   G    W   F   I   C    T   S   C   N    E   P   V   G    V   T   D

601  AAATTTGAGA ACAGGAAACA CATTGAATTT TCATTCAAC AGTTTGCAA AGTGAAATG AGCCCAGTG AGGTCAGGGA TTAGGAACT GCCCCATTGA
     TTTAAACTCT TGTCCTTTGT GTAACTTAAA AGTAAGTTG TCAAACGTT TCACTTTAC TCGGGTCAC TCCAGTCCCT AATCCTTGA CGGGGTAACT

167  K   F   E   N    R   K   H   I    E   F   S   F    Q   P   V   C    K   A   E   M    S   P   S   E    V   S   D   Q

701  ACGCCTTCCT CGCTAATTGG AACTAATTGA ATAAAAACAC CAAACCTCCT CACT
     TCCGGAAGGA GCGATTAACC TTGATTAACA TATTTTTGTG GTTTGGAGGA GTGA
```

Fig. 4


```
  1  CCAGGCCCAA  GCNTCCCCAC  CATGAGATTT  GTTCACACGA  GTCGAAAGGT  GAAGAGCTTA  AACCCGAAGA  AATTCAGCAT  TCATGAACAG  GATCACAAAG
     GGTCCGGGTT  CGNAGGGGTG  GTACTTAAAA  CAAGTGTGCT  CAGCTTTCCA  CTTCTCGAAT  TTGGGCTTCT  TTAAGTCGTA  AGTACTGGTC  CTAGTGTTTC

101  TACTGGCCTG  GACTCTGGGA  AATCTAATAGC  AGTTCCAGAT  AAAACTACA  TACGGCCAGA  GATCTTCTTT  GCATCAGCCT  CATCCTTGAG  CTCAGCCTCT
     ATGACCGGAC  CTGAGACCCT  TAGAGATATCG  TCAAGGTCTA  TTTTTGATGT  ATGCGGGTCT  CTAGAAGAAA  CGTAGTCGGA  GTAGGAACTC  GAGTCGGAGA

201  GCGGAGAAAG  GAAGTCCGAT  TCTCCTGGGG  GTCTCTAAAG  GGGAGTTTA  TCTCTACTGT  GACAAGGTA  AAGGACAAAG  TCNTCCATCC  CTTCAGCTGA
     CGCGTCTTTC  CTTCAGGCTA  AGAGGACCCC  CAGAGATTTC  CCCTCAAAAC  AGAGATGACA  CTGTTCCTAT  TTCCTGTTTC  AGTAGGTAGG  GAAGTCGACT
                ^orf                          I   L   L   G   V   S   K   G   E   F   C   L   Y   C   D   K   D   K   G   Q   S   H   P   S   L   Q   L   K
                ^84664.f1

301  AGAAGAGAA  ACTGATGAAG  CTGGCTGCCG  AAAAGGAATC  AGCACGCCGG  CCCTTCATCT  TTTATAGGGC  TCAGTGGGGC  TCGTGGAACA  TGCTGGAGTC
     TCTTCTCTT  TGACTACTTC  GACCGACGGG  TTTTCCTTAG  TCGTGCGGCC  GGGAAGTAGA  AAATATCCCG  AGTCACCCCG  AGCACCTTGT  ACGACCTCAG
     K   E   K   L   M   K   L   A   A   Q   K   E   S   A   R   R   P   F   I   F   Y   R   A   Q   V   G   S   W   N   M   L   E   S

401  GGCGGCTCAC  CCCGGATGGT  TCATCTGCAT  CTCCTGCAAT  TGTAATGAGC  CTGTTGGGGT  GACAGAGTAA  TTGGAGACA  GGAAACACAT  TGAATTTTCA
     CCGCCGAGTG  GGGCCTACCA  AGTAGACGTA  GAGGACGTTA  ACATTACTCG  GACAACCCCA  CTGTCTCATT  AACCTCTGT  CCTTTGTGTA  ACTTAAAAGT
     A   A   H   P   G   W   F   I   C   I   S   C   N   C   N   E   P   V   G   V   T   D   K   F   E   N   R   K   H   I   E   F   S

501  TTTCAAACAG  TTTGCAAAGC  TGAAATGAGC  CCANGTGAGG  TCAGCGATTA  GGBAACTGCC  CCAATGAACG  CCALTGAACG  TAATTGAAC  TAATTGTATA
     AAAGTTTGTC  AAACGTTTCG  ACTTTACTCG  GGTCACACTCC  AGTCGCTAAT  CCTTTGACGG  GGTTACTTGC  GGTAACTTTGC  ATTAAACTTG  ATTAACATAT
     F   Q   P   V   C   K   A   E   M   S   P   S   E   V   S   D   O
                                        ^84664.p1          ^84664.r1

601  AAAACCCAA  ACCTCCTCAC  TAAAAAAAA
     TTTTGGGGTT  TGGAGGAGTG  ATTTTTTTT
```

Fig 5A

```
  1  GTCGAACCAC GCGTCCGAAG CTGCGATTCAG TCCCCTGGAG TGTAGATAAG GACCCTTTCT TGCCAGGTGC TGAGACAACC ACACTATGAG
     CACGTGGGTG CGCAGCGTTC GACGACGCTG  GTGCTAAGTC AGGGACGCTG ACATCTATTT CTGGAGAAGA ACGGTCCACG TGTGATACTC
                                                                                                   M  R

     ^insert starts

  1                                                                  92929.AH1.421.Asc.f, 5'Tag: AAAGGAAAAGGGCGGCC'
                                                                                                   'MET

101  AGGCACTCCA GGAAGGCCTG ATGGTGGAGG AAGGGCGGTC TATCAATGAA TCACTGTTGC TGTTACGACA CAGAGCTCT  TGAGCAGAGC
     TCCGTGAGGT CCTTCCGGAC TACCACCTCC TTCCCGCAGG ATAGTTACTT AGTGACAACG ACAATGCTGT GTCTCGAGA  ACTCGTCTCG
 3   G  T  P   G  D  A  D   G  G  R  A   V  Y  Q  S   I  T  V  A   V  I  T   Q  R  A   E  Q  G

201  AGAGGGGATC CCATTTATTT GGGAATCCAG AATCCAGAAA TGTGTTTGTA GTTGTAGAAG AGCCCACATT GCAACGAAGA GAGCAGAAGA
     TCTCCCCTAG GGTAAATAAA CCCTTAGGTC TTAGGTCTTT ACACAAACAT CAACATCTTC TCGGGTGTAA CGTTGCTTCT CTCGTCTTCT
 36  R  G  D   P  I  Y  L   G  I  Q   N  P  E  M   C  L  Y   V  V  E   E  P  T  L   Q  R  R   E  Q  K  I

301  TCATGGATCT GTATGGCCAA CCCGAGCCCG TGAAACCCTT CCTTTTCTAC CGTGCCAAGA CTAGCACCCT GAGTGTGCCG CCTTCCCGAA
     AGTACCTAGA CATACCGGTT GGGCTCGGGC ACTTTGGGAA GGAAAAGATG GCACGGTTCT GATCGTGGGA CTCAGACGGC GGAAGGGCTT
 70  M  D  L   Y  G  Q   P  E  P  V   K  P  F   L  F  Y   R  A  K  T   S  T  L   E  S  V  A   P  D

401  CTGGTTCATT GCCTCCTCCA AGAGAGACCA GCCCATCATT CTGACTTCAG AACTGGGAAA GTCATACAAC ACTGCCCCTG AATTAAATAT AAAGTGACTGA
     GACCAAGTAA CGGAGGAGGT TCTCTCTGGT CGGGTAGTAA GACTGAAGTC TTGACCCTTT CAGTATGTTG TGACGGGGAC TTAATTTATA TTTCTGACT
103  W  F  I   A  S  S   K  R  D  Q   P  I  I   L  T  S   E  L  G  K   S  Y  N   T  A  P  E   L  N  I   N  D   O

501  ACTACACCTA GAGGTGGCAG CTTGGCTCCA CTTGGTTCCC CAATGTGTTT TCTCGTCACAT TGTCTCACAC TTTCTTGGTG TCATTTTCAC GCTCGTGCTG
     TGATGTGGAT CTCCACCGTC GAACCGAGGT GAACCAAGGG GTTACACAAA AGAGCAGTGTA ACAGAGTGTG AAAGAACCAC AGTAAAAGTG CGAGCACGAC

601  AGACAGGAGC AAGGCTGCTG TTATCATCTC ATTTTATAT GAAGAAGGAG CAATGCTTGC ATACGACTG AAGAGACAGA TGTGGCCTCA GAAGCAGGAG
     TCTGTCCTCG TTCCGACGAC AATAGTAGAG TAAAATATTA CTTCTTCCTC GTTACGAACG TATGCTGAC TTCTCTGTCT ACACCGGAGT CTTCGTCCTC

701  AGCTGGGTGG TATAAGGGTG TCCTCTCAAG CTGGTGCTGT GTAGGGCACA AGGCATCTGC ATGAGTGACT TTAAGTCTCA AAGACCAAAC ACTAAGTCTT
     TCGACCCACC ATATTCCCAC AGGAGAGTTC GACCACGACA CATCCGGTAT TCCGTAGACG TACTCACTGA AATTCAGAGT TTCTGGTTTG TGATTCAGAA
```

'92929.AH1

```
801   CTTCTAGGGG TGGGTATGAA GATCTCATCG AGCTCATGCG CCTTACCCAC CATGCCATGA CTAGCACAGA GCTGATCTCT GTTTCTGTTT TGCTTTATTC
      GAAGATCCCC ACCCATACTT CTAGAGTAGC TCGAGTACGC GCAATGGGTG GTACCGTACT GATCGTGTCT CGACTAGAGA CAAAGACAAA ACGAAATAAG

901   CCTCTTGGGA TGATATCATC CAGTCTTTAT ATGTTGCCAA TATACCTCAT TGTGTGTAT AGAACCTTCT TAGCATTAAG ACCTTGTAAA CAAAAATAAT
      GGAGAACCCT ACTATAGTAG GTCAGAAATA TACAACGGTT ATATGGAGTA ACACAGCATTA TCTGTGAAGA ATCGTAATTC TGGAACATTT GTTTTTTATTA

1001  TCTTGGGGTG GGTAATGAAGA TGCTTGAGAG CTCATGCGCG TTACCACGA TGGCATGACT AGCACAGAGC TGATCTCGT TTCTGTTTTG CTTTATTCCC
      AGAACCCCAC CCATACTTCT ACGAAGTCTC GAGTACGCGC AATGGTGCT ACCGTACTGA TCGTGTCTCG ACTAGAGACA AAGACAAAAC GAAATAAGGG

1101  TCTTGGGATG ATATCATCCA GTCTTTATAT GTTGCCAATA TACCTCATTG TGTGTATATG AACCTTCTTA GCATTAAGAC CTTGTAAACA AAAATAATTC
      AGACCCTAC TATAGTAGGT CAGAAATATA CAACGGTTAT ATGGAGTAAC ACACATTATC TGTGAAGAAT CGTAATTCTG GAACATTTGT TTTTATTAAG

1201  TCGTGTTTAAG TTAAATCATT TGTGTCCTAA ATGGAGTAAC ACACAATATC TTGGAAGAAT CGTAATTCTG GAACATTAAT ATAAAGCTAA AACTGGATATA
      AGCACAAATC AATTTAGTTA AAACAGGATT AACAGTACAC ATTAGAATTT CAATTATATT GAACACATA AATAATTAAT TATTTCGATT TTGACTATAT

1301  AAATAAAGA AGAGTAAACT G
      TTTATTTCTT TCTCATTTGA C
```

Fig. 6

```
  1  AAGCTTGCTGG  AGCCACGGATT  CAGTCCCCTG  GACTGTAGAT  AAAGACCCTT  TCTTGCCAGG  TGCTGAGACA  ACCACACTAT  GAGAGGCACT  CCAGGAGACG
     TTCCACGACC   TCGGTCGCTAA  GTCAGGGGAC  CTGACATCTA  TTTCTGGGAA  AGAACGGTCC  ACGACTCTGT  TGGTGTGATA  CTCTCCGTGA  GGTCCTCTGC

101  CTGAATGGTGG  AGGAAGGGCC   GTCTATCAAT  CAATGACTGT  TGCTGTTATC  ACATCGCAAGT ATCCAGAGGC  TCTTGAGCAA  GGCAGAGGGG  ATCCCAATTTA
     GACTACCACC   TCCTTCCCGG   CAGATAGTTA  GTTAGTGACA  ACGACAATAG  TGTACGTTCA  TAGGTCTCCG  AGAACTCGTT  CCGTCTCCCC  TAGGGTAAAT

201  TTTGGGAATC   CAGAATCCAG   AAATGTGTTT  GTATTGTGAG  AAGGTGGA
     AAACCCTTAG   GTCTTAGGTC   TTTACACAA   CATAACACTC  TTCCAACCT
```

Fig 7

```
1    ATGGTCCTGA GTGGGGCGGCT GTGCTTCCGA ATGAAGGACT CGGCATTGAA GGTGCTTTAT CTGCATAATA ACCAGCTTCT AGCTGGAGGG CTGCATGCAG
     TACCAGGACT CACCCCGCGA CACGAAGGCT TACTTCCTGA GCCGTAACTT CCAGAAATA  TGGTCGAAGA TCGACCTTCC GACGTACGTC
1    M  V  L  S  G  A  L  C  F  R  M  K  D  S  A  L  K  V  L  Y  L  H  N  N  Q  L  L  A  G  G  L  H  A  G

101  GGAAGGTCAT TAAAAGGTGAA GAGATCAGCG TGGTCCCCAA TCGGTGGCTG GATGCCAGCC CTGTCATCCTG GTCCAGGGTG GAAGCCAGTG
     CCTTCCAGTA ATTTCCACTT CTCTAGTCGC ACCAGGGGTT AGCCACCGAC CTACGGTCGG ACAGGTCCCA  CAGGTCCCAC CTTCGGTCAC
35   K  V  I  K  G  E  E  I  S  V  V  P  N  R  W  L  D  A  S  P  V  I  L  G  V  Q  G  G  S  Q  C

201  CCTGTCATGT GGGGTGGGGAC AGGAGCCGAC TCTAACACTA GAGCCAGTGA ACATCATGGA GCTCTATCTT GGTGCCAAGAG CTTCACCTTC
     GGACAGTACA CCCCACCCCG TCCTCGGCTG AGATTGTGAT CTCGGTCACT TGTAGTACCT CGAGATAGAA CCACGGTTCC GAAGTGGAAG
68   L  S  C  G  V  G  Q  E  P  T  L  T  L  E  P  V  N  I  M  E  L  Y  L  G  A  K  E  S  K  S  F  T  F

301  TACCGGCGGG ACATGGGGCT CACCTCCAGC TTCGAGTCGG CTGCCTACCC GGGCTGGTTC CTGTGCACGG TCCCTGAAGC CGATCAGCCT GTCAGACTCA
     ATGGCCGCCC TGTACCCCGA GTGGAGGTCG AAGCTCAGCC GACGGATGGG CCCGACCAAG GACACGTGCC AGGGACTTCG GCTAGTCGGA CAGTCTGAGT
101  Y  R  R  D  M  G  L  T  S  S  F  E  S  A  A  Y  P  G  W  F  L  C  T  V  P  E  A  D  Q  P  V  R  L  T

401  CCCAGCTTCC CGGAGAATGT GGCTGGAAATG AGACTCTAC TTCCAGCAGT GTGACTAG
     GGGTCGAAGG GCTCTTACCA CCGACCTTAC TCTGAGATG AAGGTCGTCA CACTGATC
135  Q  L  P  E  N  G  W  N  A  P  I  T  D  F  Y  Q  Q  C  D  O
```

Fig. 8

```
  1  GCTCCCGCCA GGAGAAAGGA ACATTCTGAG GGGAGTCTAC ACCCTGTGGA GCTCAAGATG GTCCTGAGTG GGGCGGCTGTG CTTCCGAATG AAGGACTCGG
     CGAGGGCGGT CCTCTTTCCT TGTAAGACTC CCCTCAGATG TGGGACACCT CGAGTTCTAC CAGGACTCAC CCCGGCGCAC GAAGGCTTAC TTCCTGAGCC
  1   I  A  P  A  R  K  E  H  S  E  G  S  L  H  P  V  E  L  K  M  V  L  S  G  A  L  C  F  R  M  K  D  S  A

101  CATTTCGAGGT GCTTTATCTG CATAATAACC AGCTTGTAGC TGGAGGGCTG CATGCAGGGA AAGTCATTAA AGGTGAAGAG ATCAGCGTGG TCCCCAATCG
     GTAAACTTCCA CGAAATAGAC GTATTATTGG TCGAACATCG ACCTCCCGAC GTACGTCCCT TCAGTAATT TCCACTTCTC TAGTCGCACC AGGGGTTAGC
 35   L  K  V  L  Y  L  H  N  N  Q  L  L  A  G  G  L  H  A  G  K  V  I  K  G  E  E  I  S  V  V  P  N  R

201  GTGGCTGGAT GCCAGCCTGT CCCCGTCGTC CCTGGGTGTC CAGGGTGGAA GCCAGTGCCT GTCATGTGGA GTGGGGCCAG AGNGGACTCT AACAT
     CACCGACCTA CGGTCGGACA GGGGCAGCAG GGACCCACAG GTCCCACCTT CGGTCACGGA CAGTACACCT CACCCCGGTC TCNCCTGAGA TTGTA
 68   W  L  D  A  S  L  S  P  V  I  L  G  V  Q  G  G  S  Q  C  L  S  C  G  V  G  Q  E  X  T  L  T
```

Fig 9A

```
  1  ATAGGGAATT TGGCCCTCGA GGCCAAGAAT TCGGCACGAG GGGAGCCTGC TTTCTACTTA GGTCTCAAAT TTTCCAGCCT TGTCTTTGCC TAAAATTTCC
  1  TATCCCTTTA ACCGGGAGCT CCGGTTCTTA AGCCGTGCTC CCCTCGGACG AAAGATGAAT CCAGATTTTA AAAGGTCGGA ACAGAAACGG ATTTTAAAGG
                                                      ^insert starts

101  TGCTGTTTAT TTCAAAATAG GGTCTACATA CTGTGGAGCT CATGATGGTT CTGAGTCGGGA CACTATGCTT CCGAATGAAG GATTCAGCCT TGAAGGTACT
101  ACGACAAATA AAGTTTTATC CCAGATGTAT GACACCTCGA GTACTACCAA GACTCACCCC GTGATACGGAA GGCTTACTTC CTAAGTCGGA ACTTCCATGA
                                                 M  V  L  S  G  A  L  C  F  R  M  K  D  S  A  L  K  V  L
                                                 'orf

201  GTATCTGCAC AATAACCAGC TGCTGGGCTGG AGGACTGCAC GCGAGAGAAG TCATTAAAGG TGAAGAGATC AGTGTGGTGCC CAAATCGGGC ACTGGATGCC
 20  CATAGACGTG TTATTGGTCG ACGACCCGAC TCCTGACGTG CGCTCTCTTC AGTAATTTCC ACTTCTCTAG TCACAACAGG GTTTAGCCCG TGACCTACGG
     Y  L  H   N  N  Q  L  A  G  G  L  H  A  E  K  V  I  K  G  E  I  S  V  V  P  N  R  A  L  D  A

301  AGTGTGTCCC CTGTCATCCT GGGCGTTCAA GGAGGAAGCC AGTGCCTATC TTGTGGGACA GAGAAAGGGC CAATTGTGAA ACTTGAACCA GTGAACATCA
 53  TCAGACAGGG GACAGTAGGA CCCGCAAGTT CCTCCTTCGG TCACGGATAG AACACCCTGT CTCTTTCCCG GTTAACACTT TGAACTTGGT CACTTGTAGT
     S  L  S  P  V  I  L  G  V  Q  G  G  S  Q  C  L  S  C  G  T  E  K  G  P  I  L  K  L  E  P  V  N  I  M

401  TGGGACTCTA CCTCGGGGGC AAGAATCAC AGACTCAC CTTCTACCGG CGGGATATGG GTCTTCCTTC CACCTTCGAA TCCGCTGGCT ACCCAGGGTG
 87  ACCTGAGAT GGAGCCCCCG TTCCTTAGTT TCTGAGAGTG GAAGATGGCC GCCCTATACC CAGAATGGAAG GTGGAAGCTT AGGCAGCGGA TGGGTCCCAC
     E  L  Y   L  G  A  K  E  S  K  S  F  T  F  Y  R  R  D  M  G  L  T  S  S  F  E  S  A  A  Y  P  G  W

501  GTTCCTCTGC ACCTCACCCG AAGCTGACCA GCCTGTGAGG CTACTACAGA TCCTTAGGA CCCGGCTGG GATGCTCCGA TCACAGACTT CTACTTCAG
120  CAAGGAGACG TGGAGTGGGC TTCGACTGGT CGGACACTCT GGATGATGTCT AGGAATTCCT GGGCGGACC CTACGAGGCT AGTGTCTGAA GATGAAAGTC
     F  L  C   T  S  P  E  A  D  Q  P  V  R  L  T  Q  I  D  P  A  W  D  A  P  I  T  D  F  Y  F  Q

601  CAGTTGTAGT AGGGCTCCGT GGTCCCGAA ACTCCATAAG CAGAGGCAGA GTAGGCAGTG GCGGCTCCTG ATAGAGGATA GAGAGACGA GGAGCTCAC
153  GTCACACTGA TCCCGAGGCA CCAGGGCTT TGAGGTATC GTCTCCGTCT CATCCGTGAC CGCCGAGGAC TATCCTCTAT CTCTCTGCT CCTCGAGGTG
     Q  C  D   D

701  AGTAGCTGGC TTACTCCTCT CCTTCCTAC TGGACTCCCG CTTCTGACCT AAGGCACACA GACACTCTCT TCTCCTCGAT CCCAGTGCTG GTAAAACTTC
701  TCATCGACCG AATAGGAGGA GGAAGGGATG ACCTGAGGGC GAAGACTGGA TTCCGTGTGT CTGTGAGAGA AGAGGAGCTA GGGTCACGAC CATTTTGAAG
```

```
 801  TGGTATTTCG  AGCTCAATCT  GTGATTCTTT  TCAGATTGGA  TGTGACTACC  TCTGCGTCGG  AACCCGAATAG  AAACCACGTA  GGACCAACAA  AGAGCAACAT
      ACCATAAACC  TCGAGTTACA  CATCTAAGAA  AGTCTAACCT  ACCATGATGG  AGACCACACC  TTGGGTTATC  TTTGGTGCAT  CCTGGTTGTT  TCTCGTTGTA

 901  AAAAGATTCT  TGGGTGAAGA  AGAGGTGGGA  ACTGTGTCATA CATAGTAAGA  TCTGACACAG  TACCTGGGAA  GTCCTGGCCAT TCCTTATGTT  CTGGAGAAAG
      TTTTCTAAGA  ACCCACTTCT  TCTCCACCCT  TGACAGTAT   GTATCATTCT  AGACTGTGTC  CAGGACGGTA  AGGAATACAA  GACCTCTTTC

1001  TGGAGGGGGG  GTCACCAAGA  CTTTCTCTGG  CTGGCTGGGC  CCTTTCCCTC  ACATCTGCAG  CCTCTCTCAT  TCTTGCCTTC  AATTCTCTGGC
      ACCTCCCCCC  CAGTGGTTCT  GAAAGAGACC  GACCGACCCG  GGAAAGGAG   TGTAGACGTC  GGAAGAGTA   AGAACGGAAG  TAAGAGACCG

1101  CCTGAACCGA  GAGGGTGATA  TCAGGATAAC  TGACAGAGGA  CACTGTCCTG  GTTTGAAACC  AGAGGGACA   ATAAAAAACC  CTGATTCTGG
      GGACTTGGCT  CTCCCACTAT  AGTCCTATCG  ACTGTCTTCT  ACTGGTCCGT  GTGACAGGAC  CAAACTTTGG  TCTCCCCTGT  TATTTTTTGG  GACTAAGACC

1201  TCTCCTACTCA CATAAAAAGA  AGCTTGTTGA  CATTAAGTGG  GAAGAGATTG  CTACTAAATA  ACATTACCTTG TAATTTCATC  TTAATTAAAA  TATACTCTC
      AGAGATGAGT  GTATTTTTCT  TCGAACACTT  GTAATTCACC  CTTTCTCTAAC GATGATTTAT  TGTATGAC    ATTAAAGTAG  AATTAATTTT  ATATGAAGAG

1301  TATATATTAT  ATTTTAAAA   AAAAAAAAA   AAAAAAAAAA  AAAAACATGC  GGCCGCGAAGC TTAATTCCATT TAGGA
      ATATAATATA TAAAATTTTT  TTTTTTTTTT  TTTTTTTTT   TTTTTTGTACG CCGGCGCTTCG AATTAAGGTAA ATCCT
```

`insert ends`

Fig. 9B

Fig. 10

```
1    GCGAGCCTGCT TTCTACTTAG GTCTCAAATT TTCCAGCCTT GTGTTTGCCT AAAATTCCTT TCAAAATAGG GTCTACATAC TGTGGAGCTC
     CCTGGGACGA AAGATGAATC CAGAGTTTAA AAGGTCGGAA CAGAAACGGA TTTTAAAGGA AGTTTTATCC CAGATGTATG ACACCTCGAG

101  ATGATGGTTC TGAGTGGGGC ACTATGCTTC CGAATCGAAG GAAGGTACTG TATTCGCACA ATAACCAGCT GCTGGCTGGA GGACTCGCAG
     TACTACCAAG ACTCACCCCG TGATACGAAG GCTTAGCTTC CTTCCATGAC ATAAGCGTGT TATTGGTCGA CGACCGACCT CCTGAGCGTC
1    M   M   V   L   S   G   A   L   C   F   R   M   K   D   S   A   L   K   V   L   Y   L   H   N   Q   L   L   A   G   G   L   H   A
~orf

201  CAGAGAAGGT CATTAAAGGT GAGGAGATCA GTGTTGTCCC AAATGGGCCA CTGGATGCCA GTCTGTCCCC TGTCATCCTG GGCGTTCAG GAGGAAGGCA
     GTCTCTTCCA GTAATTTCCA CTCCTCTAGT CACAACAGGG TTTACCCGGT GACCTACGGT CAGACAGGGG ACAGTAGGAC CCGCAAGTC CTCCTTCCGT
35   E   K   V   I   K   G   E   E   I   S   V   V   P   N   R   A   L   D   A   S   L   S   P   V   I   L   G   V   Q   G   G   S   Q

301  GTGCCTATCT TGTGGGACAG AGAAAGGGCC AATTCTGAAA CTTGAGCCAG TGAACATCAT GGAGCTCTAC CTCGGGGCCA AG
     CACGGATAGA ACACCCTGTC TCTTTCCCGG TTAAGACTTT GAACTCGGTC ACTTGTAGTA CCTCGAGATG GAGCCCCGGT TC
68   C   L   S   C   G   T   E   K   G   P   I   L   K   L   E   P   V   N   I   M   E   L   Y   L   G   A   K
```
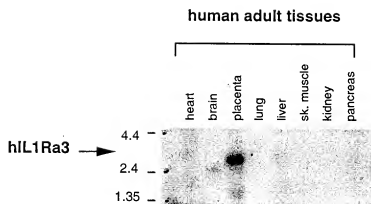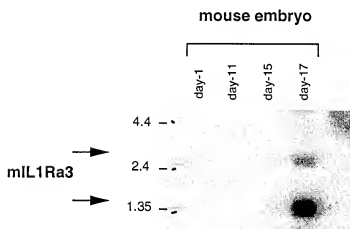
A human adult tissues

heart, brain, placenta, lung, liver, sk. muscle, kidney, pancreas

hIL1Ra3 →

4.4 –
2.4 –
1.35 –

B mouse embryo

day-1, day-11, day-15, day-17

4.4 –
2.4 –
mIL1Ra3 →
1.35 –

Fig. 11

```
hIL1Ra      1  M E I C R G L R S   I T L L L F L F H S E T I C R P S G R K S S K M Q A _ I W D V N Q K T F Y L
hIL1Rabeta  1  · · · · · · · · · · · · · M R G T P G D A D G G G R A V Y Q S M C K P I T G T I N D L N Q Q V W T L
hIL1Ra1     1  · · · · · · · · · · · · · · · · · · · · · · · · C R G P K V K N L N P K K F S I H D Q D H K V L V L
hIL1Ra2     1  · · · · · · · · · · · · · M R G T P G D A D G G G · · · · · · · · · · · · · · · · · · · · · · · · · ·
hIL1Ra3     1  · · · · · · · · · · · · · · · · · · · · · · · · · · M V L S G A L C F R M K D S A L K V L Y L
mIL1Ra3     1  · · · · · · · · · · · · · · · · · · · · · · · · · · M V L S G A L C F R M K D S A L K V L Y L

hIL1Ra     51  R N N Q L V A G Y L Q G P · N V N L E E K I D V V P · · · · I E P H A L F L G I H G G K M C L S
hIL1Rabeta 38  Q G G N L V A V P R S D S V T P V T V A V I T C K Y P E A L E Q G R G D P I Y L G I Q N P E M C L Y
hIL1Ra1    27  D S G N L I A V P D K N · Y I R P E I F F A L A S S L S S A S A E K G S P I L L G V S K G E F C L Y
hIL1Ra2    13  · · · · · R A V Y Q S · · · I T V A V I T C K Y P E A L E Q G R G D P I Y L G I Q N P E M C L Y
hIL1Ra3    22  H N N Q L L A G G L H A G · · K V I K G E E I S V V P N R W L D A S L S P V I L G V Q G G S Q C L S
mIL1Ra3    22  H N N Q L L A G G L H A E · · K V I K G E E I S V V P N R A L D A S L S P V I L G V Q G G S Q C L S

hIL1Ra     94  C V K S G · · D E T R L Q L E A V N I T D L S E N R K · Q D K R F A F I R S D S G P T S F E S A A
hIL1Rabeta 88  C E K V G · · E Q P T L Q L K E Q K I M D L Y G Q P E · P V K P F L F Y R A K T G R T S T L E S V A
hIL1Ra1    76  C D K D K G Q S H P S L Q L K K E K L M K L A A Q K E S A R R P F I F Y R A Q V G S W N M L E S A A
hIL1Ra2    53  C E K V G · · E Q P T L Q L K E Q K I M D L Y G Q P E · P V K P F L F Y R A K T G R T S T L E S V A
hIL1Ra3    70  C G V G · · · Q E P T L T L E P V N I M E L Y L G A K · E S K S F T F Y R R D M G L T S S F E S A A
mIL1Ra3    70  C G T E · · · K G P I L K L E P V N I M E L Y L G A K · E S K S F T F Y R R D M G L T S S F E S A A

hIL1Ra    141  C P G W F L C T A M E A D Q P V S L T N M P D E G · · · V M V T K F Y F Q E D E · · · · · ·
hIL1Rabeta 135 F P D W F I A S S K R D Q P I I L T S E L G K S Y N · · · T A F E L N I N D · · · · · · · · ·
hIL1Ra1   126  H P G W F I C T S C N C N E P V G V T D K F E N R K H I E F S F Q P V C K A E M S P S E V S D
hIL1Ra2   100  F P D W F I A S S K R D Q P I I L T S E L G K S Y N · · · T A F E L N I N D · · · · · · · · ·
hIL1Ra3   116  Y P G W F L C T V P E A D Q P V R L T Q L P E N G G W N A P I T D F Y F Q Q C D · · · · · ·
mIL1Ra3   116  Y P G W F L C T S P E A D Q P V R L T Q I P E D P A W D A P I T D F Y F Q Q C D · · · · · ·
```
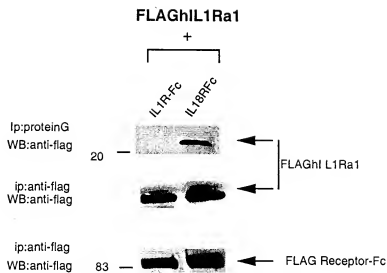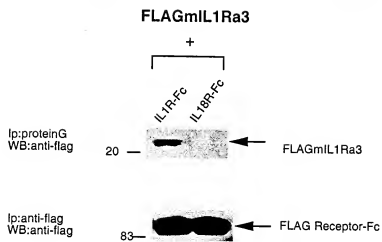
Fig. 12.

FLAGhIL1Ra1

+

IL1R-Fc   IL18RFc

Ip:proteinG
WB:anti-flag   20 —   FLAGhI L1Ra1

ip:anti-flag
WB:anti-flag

ip:anti-flag
WB:anti-flag   83 —   ◄ FLAG Receptor-Fc

Fig.13

FLAGmIL1Ra3

+

IL1R-Fc   IL18R-Fc

Ip:proteinG
WB:anti-flag   20 —   ← FLAGmIL1Ra3

Ip:anti-flag
WB:anti-flag   83—   ← FLAG Receptor-Fc

Fig 14